Two Sided Matching: An Algorithm with both sides making proposals

Aditya V Kuvalekar *

Abstract

The most commonly used algorithm in a classical two-sided marriage market is the one due to Gale-Shapley. ([3]). The algorithm always obtains a stable match. However, the outcome is the most preferred stable match for the side proposing and also the least preferred stable match for the side receiving the proposals. I propose a new algorithm in the classical two-sided marriage market. The algorithm has an appealing feature that both sides of the market make proposals in every round. The algorithm does not distinguish (ex-ante) between men and women at any stage. The ex-post distinction arises primarily because in every round potential cycles are formed that are broken randomly. The algorithm always produces a stable matching on all profiles and has appealing fairness properties. Lastly, the algorithm frequently obtains a different stable matching than the *Men Optimal* or the *Women Optimal* in many markets.

1 Introduction

I study the classical two-sided marriage market due to Gale-Shpley ([3]). In this market there are two sides - Men and Women and they have strict preferences over the other side. A matching is an assignment wherein each man is matched to a woman or himself (indicating that he is unmatched) and vice-versa. Moreover, no two men are assigned a same woman. A desirable property in the context of two-sided marriage market is *stability.* Stability requires that no subset of men and women would want to deviate and form their own assignment. The question about the existence of a stable matching in marriage market was settled positively in a seminal paper by D.Gale and L.Shapley in 1962 ([3]) by proposing a *Deferred Acceptance Algorithm* (DAA henceforth). The algorithm established that in a two-sided marriage market with strict preferences, there always exists at least one stable matching. Moreover, the authors also prove that the matching obtained through DAA is the best matching for the proposing side amongst the set of stable matchings.

*I am indebted to Jay Sethuraman for countless discussions and guidance in the development of this algorithm. I sincerely thank Efe Ok and Ennio Stacchetti for their useful comments.

Following this work, Knuth (1976)([7]) proved that when all agents have strict preferences the common preferences of two sides of the market are opposed to each other on the set of stable matchings. That is, for any two stable matchings, μ_1 and μ_2, all the men **weakly** prefer μ_1 to μ_2 *if and only if* all the women weakly prefer μ_2 to μ_1. As a natural corollary, we obtain that any DAA outcome is not only the best for the proposing side but it is also the worst for the side receiving proposals. This is slightly unsatisfactory that the only algorithms we have so far should give us only two extremes and for getting anything else we should have to compute a linear program. This raises a questions as to can we devise a procedure that does not distinguish between two sides explicitly and gives a stable match. With this focus, I propose a new algorithm wherein both sides of the market propose in each round to a set of agents. The proposed algorithm yields a stable matching whenever it terminates. I have examples where it does not terminate i.e. it cycles. Based on those I have a resolution in case the algoritm cycles.

The most notable feature of the algorithm are that it is the first algorithm wherein both sides of the market propose together. The manner in which they do so has flavours of a decentralized market and therefore seem natural. Moreover, the algorithm does not distinguish between men and women at any stage. Neither does the algorithm give preference to a particular agent ex-ante.

Another strand of literature where this algorithm fits naturally is the literature related to fairness in the context of stable matching. To quote Donald Knuth ,

> The different algorithms considered until now favour the men, and if we interhange the roles of men and women they would become favorable to the women. Such injustice is too shocking for the present day. Can we therefore find a solution that treats both sexes fairly?
>
> We can enumerate all the stable solutions and choose the most satisfying matching according to certain criteria. This might take a long time if the number of stable matchings is large. We do not know in general if there exists a large or small number of solutions. If therefore seems preferable to use another method.

The primary motivation of the algorithm is a procedure that treats both sexes equally. It is well-known that the set of stable matching has a specific lattice structure (Roth and Sotomayor, 1990 [10]). and the two Gale-Shapley procedures pick the two extreme points of this lattice. There have been papers trying to address this issue of procedures picking the only extreme points. In the world of ordinal preferences such as in Two-Sided matching, giving a structured theoretical foundation to the question of fairness is challenging. Masarani and Gokturk (1990)[9] showed impossibilities to obtain a fair stable matching wherein their notion of *fairness* was Rawlsian. While negative, this result demanded that the rule be deterministic. To circumvent this problem one way to recover fairness is to use probabilistic (stable) matching mechanisms that are ex ante fair and/or 'procedurally fair', as in Klaus and Klijn (2006)[5], and Ma (1996)[8].

Ma(1996)[8] proposed a procedure wherein we start with a *Random Priority* over agents. Following the priority, we start with an empty match and add agents one by one to the matching satisfying all the blocking pairs *within the match*. That is, blocking pairs outside the agents not added so far are ignored. While this procedure does not distinguish between sexes or any particular agent, the support of the distribution obtained through this procedure is the entire set of stable matching. In contrast, the algorithm I propose has a narrower support and in many examples (such as those in Subsection 3.1 on the motivation of the algorithm) it will pick a strict subset of stable matchings that do not involve any of the extreme outcomes due to Gale-Shapley.

Using another approach, Teo and Sethuraman (1998)[13] and Sethuraman et al.(2004)[12] established the existence of a deterministic 'compromising mechanisms' for marriage and college admissions models respectively. Specifically, they showed that if all agents order their (possibly non-distinct) matches at the, say, k stable matchings from best to worst, then the map that assigns to each agent of one side of the market its l-th best match and to each agent of the other side its $(kl+1)^{st}$ best match constitutes a stable matching. Teo and Sethuraman (1998)[13] and Sethuraman et al. (2004)[12] used linear programming tools to prove that these '(generalized) median stable matchings' are indeed well-defined and stable. The main issue with the *median stable matching* is that to obtain those we have to compute the entire set of stable matchings. Secondly, and perhaps more importantloy, there is no axiomatic foundation for the *Median Stable Matching* rule.

Lastly, in an experiment on two-sided one-to-one matching market, Echenique and Yariv [2] found that median matchings were selected more often than any other matching. The setting they had was decentralized in nature. The most striking departure from the standard setting in this case is that unlike in Gale-Shapley, here both sides could make proposals. Therefore, it is not entirely surprising that the resulting matching obtained was different than the two extreme Gale-Shapley outcomes. Given the decentralized nature of the algorithm wherein both sides of the market propose, it is reasonable to expect that a procedure like the one proposed would be prevalent in markets that do not have a centralized clearinghouse.

The paper is organized as follows. In section 2, we discuss the preliminaries of the model and the setup. This is a standard setup of the classical two-sided marriage market and can be skipped by those familiar with the literature. Section 3 describes the algorithm in detail. Section 4 has a number of examples illustrating the execution of the algorithm along with an example of an instance where the algorithm cycles. Section 5 is devoted to the proofs of the fact that the algorithm always yields a stable matching. Section 6 describes the fairness properties of the algorithm. And the last section (7) is devoted to the conclusion.

2 Preliminaries

2.1 Setup

There are two finite disjoint sets of agents, call them *Men* and *Women*. $M = \{m_1, m_2, \ldots, m_l\}$ is a set of men.

$W = \{w_1, w_2, \ldots, w_k\}$ is a set of women.

$N := M \cup W$. We denote a generic agent by i. If we need to distinguish between a man and a woman for illustration, we will do so by labeling them as m_j and w_l respectively. The structure we will use henceforth is of the form where $N = \{1, 2, \ldots, n\}$ where $n = |M| + |W|$. Agents from 1 to $|M|$ are men and $|M| + 1$ to n are women. This way, we will not need to refer to men and women separately for most of the proofs.

Each agent has a linear ordering over the agents on the other side of the market. Moreover, in that ordering one element is the agent himself indicating that the agents is also potentially willing to remain single over being matched to some agents from the other side. In what follows though, we will study a model wherein the agent will always prefer to be matched over remaining single. I do not believe that any results are altered in any way by relaxing this assumption. However, that needs to be proved in general.

Man m's preferences can be represented as $\succeq_m$. For a generic agent, the ordering will be denoted by $\succ_i$.

For example, $\succ_i = (1, 3, 4, 7, 5, 8, 2, 6)$ means $1 \succ_i 3 \succ_i 4 \succ_i 7 ... \succ_i 2 \succ_i 6$.

In general, to avoid distinguishing between men and women in terms of notation, we will denote by P_i the preferences of agent i over the other agents.

$$P := (P_i)_{i \in N}$$

P_i^j denotes the j^{th} ranked element according to the preferences of i.

Let $\mathcal{P}$ denote the class of all preference profiles.

Lastly, let A_i denote the set of acceptable agents for each i. That is, i would prefer being matched to any agent in i over being single, and would prefer being single than to match with any agent not in A_i.

2.2 Matching

A marriage market is a triple (M, W, P). A matching is a one-one map $\mu : N \to N$ such that $\mu(m) \in W \cup \{m\}$ $\forall$ m and $\mu(w) \in W \cup \{w\}$ $\forall$ w. Moreover, for the matching to be meaningful we also require that $\mu(m) = w$ iff $\mu(w) = m$. In general, $\mu(i) = j \Leftrightarrow \mu(j) = i$.

Let $\mathcal{M}$ be the set of all feasible matchings.

2.3 Stability of a matching

A desirable property for a matching is that it be stable.

Definition 2.1. *A matching μ is stable if $\nexists$ (i,j) such that $j \succ_i \mu(i)$ and $i \succ_j \mu(j)$. Moreover, such pair(s) (i,j) are called as blocking pair(s) in a matching μ.*

Through their algorithm Gale-Shapley (1962) established that in every marriage market there always exists a stable matching. Let $S(P)$ denote a set of stable matchings for the market (M, W, P).

2.4 Lattice Structure

Define an ordering $\succ_M$ over $S(P)$ as follows -

$$\mu \succ_M \hat{\mu}(\neq \mu) \Leftrightarrow \ \mu(m) \succeq_m \hat{\mu}_m, \ \forall m \in M.$$

Using this ordering, we can define $\lambda := \mu \vee_M \mu'$ as a function on N that assigns to each man his most preferred match from μ and μ'. Formally,

$$\lambda(m) = \begin{cases} \mu(m) & \text{if } \mu(m) \succ_m \mu'(m) \\ \mu'(m) & \text{otherwise} \end{cases} \tag{2.1}$$

$$\lambda(w) = \begin{cases} \mu(w) & \text{if } \mu'(w) \succ_m \mu(w) \\ \mu'(w) & \text{otherwise} \end{cases} \tag{2.2}$$

$\mu \wedge_M \mu'$ are also defined analogously by assigning each man his worst preferred mate amongst μ and μ' and converse for the women.

The following theorem due to Conway established that the set of stable matchings is a lattice. In fact, it is a distributive lattice.

Theorem 2.1. *If $\mu, \mu' \in S(P)$ then, $\mu \vee_M \mu', \mu \wedge_M \mu' \in S(P)$.*

Since the set of stable matchings is finite (obviously) there exists a $\succ_M$ **maximum** element. Simiarly, if we define $\succ_W$ analogously, then there exists a $\succ_W$ **maximum** element according to the order of women. Gale-Shapley proved that the version of their algorithm in which men propose yields a maximum element according $\succ_M$ in $S(P)$. Similarly, a woman proposing version of the algorithm yields an matching that is the most preferred by women.

3 Algorithm

3.1 Motivation

The primary motivation of the algorithm is to think of a procedure wherein both sides of the market make proposals.

3.1.1 Example 1

A simple example of such a market wherein one could expect to obtain an outcome different than the Gale-Shapley outcome should the two sides be allowed to make proposals is the following.

1	2	3
4	5	6
5	6	4
6	4	5

Here, $1, 2, 3$ can be thought of as Men and $4, 5, 6$ can be thought of as Women. The each column gives the preferences of a corresponding agent. That is, $4 \succ_1 5 \succ_1 6$, $5 \succ_2 6 \succ_2 4$ and so on.

Women's preferences are as below.

4	5	6
2	3	1
3	1	2
1	2	3

Observe that the outcome of the Men Proposing Gale-Shapley algorithm is $(1, 4), (2, 5), (3, 6)$. Simlarly, the women proposing Gale-Shapley yields, $(1, 6), (2, 4), (3.5)$. However, imagine a procedure wherein both sides make proposals each round to a set of agents starting from their top agents. In each round they expand their sets of acceptable agents according to their preferences. And we match agents only when they both list each other as mutually acceptable.

Then, in the first round, 1 would propose to 4, 2 to 5, 3 to 6, 4 to 2, 5 to 3 and 6 to 1. None of the agents find each other mutually acceptable. But in the second round, 1 proposes to 4,5. 2 proposes to 5,6 and so on. In this round, 1 and 5 find each other mutually acceptable. So do 2 and 6, and 3 and 4. So, we match $(1, 5), (2, 6), (3, 4)$.

3.1.2 Example 2

To further expand on the idea of both sides making proposals, consider the following preferences of men.

1	2	3	4
5	6	7	8
6	7	8	5
7	8	5	6
8	5	6	7

Women's preferences are -

5	6	7	8
2	3	4	1
3	4	1	2
4	1	2	3
1	2	3	4

The Gale-Shapley outcomes are -

$$\{(1,5),(2,6),(3,7),(4,8)\} \text{ and } \{(1,8),(2,5),(3,6),(4,7)\}.$$

We will look at this example again more formally but as of now let us look at it informally. Imagine a procedure wherein the agents from both sides of the market make proposals to a set of agents every round. If there is no agent whom they are proposing to and is also proposing them then they expand their *proposing to* set of agents in the following round. They start by proposing to the top agent according to their preferences and expand the set of agents they propose to according to their preferences.

In the first round, all the agents propose to their top choice and no pair finds each other mutually acceptable as if an agent i proposes to j in the first round then j is the top agent according to i and i is the worst agent according to j. In the following round, for example, 1 proposes to 5 and 6. On the other hand, 5 proposes to 2 and 3, while 6 proposes to 3 and 4. 2 proposes to 6 and 7. 6 proposes to 3 and 4, while 7 proposes to 4 and 1. It can be seen that there are no two agents who propose to each other in this round too.

However, in the following round, that is round 3, consider the example of agent 1. He proposes to 5,6 and 7. Observe that 6 proposes to 3, 4 and 1, while 7 proposes to 4, 1 and 2. That is, now all of a sudden there are two agents who 1 is interested in being matched to and they too are interested in being matched with 1. Those two agents are 6 and 7. 1 would prefer being matched to 6 over 7. But observe that 6 too has two mutually acceptable agents and she would prefer being matched to 4 rather than 1. Continuing this way, if we look at the top agent for each person from the list of their mutually acceptable agents in this round, we obtain a cycle as below.

$$1 \to 6 \to 4 \to 5 \to 3 \to 8 \to 2 \to 7 \to 1$$

Now, if we break this cycle randomly either in favour of men or women we obtain the following two matches -

$$\{(1,6),(2,7),(3,8),(4,5)\} \text{ or } \{(1,7),(2,8),(3,5),(4,6)\}$$

The entire algorithm described below makes this simple idea from the above two examples formal and attempts to solve a complicated cycling problem that can often arise in markets with arbitrary preferences. But at the core, the motivation and the essence is these rather simple examples.

3.2 Formal Algorithm

- Throughout the algorithm $\mu(i) = 0$ indicates that agent i is unmatched.
- Set $\mu(i) = 0 \ \forall i \in M \cup W$.
- **Round 1:**

$$S_i^1 = P_i^1, \forall i \in M \cup W \tag{3.1}$$

  Each agent applies to an agent in S_i^1. If $\exists(i, j)$ such that $i \in S_j^1$ and $j \in S_i^1$ then (i, j) are *tentatively* matched. (As will be seen later, (i, j) will never be broken.).

  That is,

$$\mu(i) = j \text{ and } \mu(j) = i$$

  In simple terms, in the first round, each agent proposes to their top choice. If it so happens that i's top choice is j and vice-versa then (i, j) are matched.

  Define, $\forall i$ such that $\mu(i) = 0$,

  Define,

$$\beta_i := 1, \forall i$$

  The purpose of defining β_i is that at some later stage, say in round 10, agent i gets his 5^{th} preferred match, say j, then we can shrink the set of agents that i looks at thereafter to his top 5. However, at some stage after that, if j leaves i for someone else then i would start proposing from his 10^{th} round set of proposals. β_i will help us keep a memory of that.

  Define,

$$\beta_i = \begin{cases} 1 & \text{if } \mu(i) > 0 \\ 2 & \text{otherwise} \end{cases} \tag{3.2}$$

  For any round k, $\beta_i \leq |A_i|$.

  Define, $\mathcal{A}$ to be the set of all feasible β's.

  Let $\mu^2 := \mu$.

  For the next round, the only relevant variables are μ_2 and β.

- **Round k** $(k \geq 2)$

  At the end of the previous round, we generate μ^k and β for the next round. That is the input for round k.

  Define each round's operator as $\mathbb{H}$.

  Define,

$$S_i^k := \begin{cases} \{j : j \succ_i \mu^k(i)\} & \text{if } \mu^k(i) > 0 \\ \{ \text{ Top } \beta_i \text{ elements according to } \succ_i\} & \text{if } \mu(i) = 0 \end{cases} \tag{3.3}$$

Define,

$$S^k := (S_i^k)_{i=1:N} \tag{3.4}$$

Definition 3.1. *The pair* (μ^k, β) *is called the Round-Input for Round* k*.*

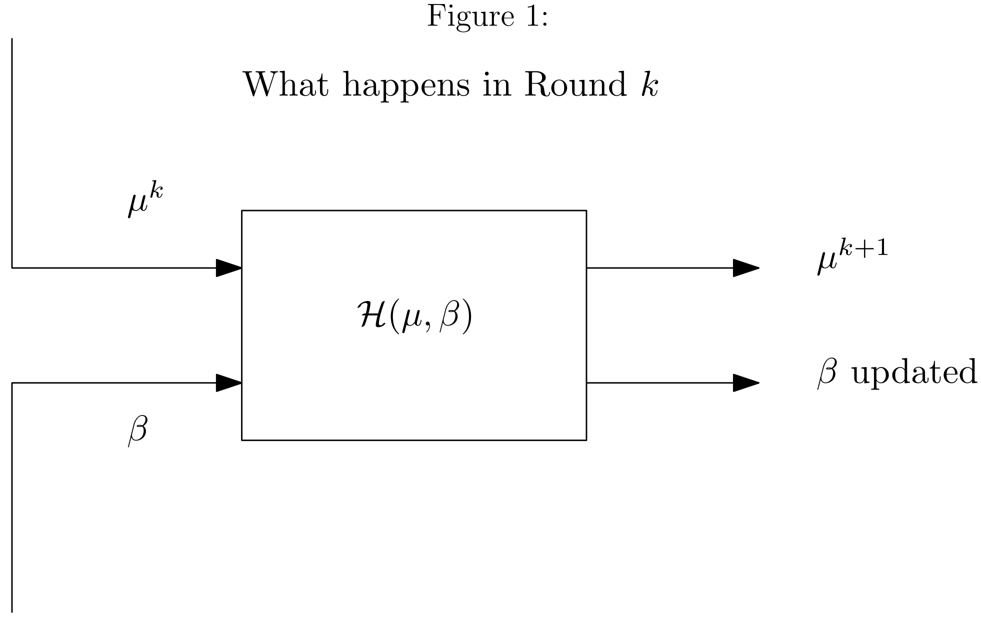
For any round k, $\mathbb{H}$ is a function that takes μ^k and β as input and outputs $\hat{\mu}^k$ and $\hat{\beta}$.

As defined earlier, $\mathcal{M}$ is a set of all feasible matchings and $\mathcal{A}$ is the set of all feasible β's.

That is, $\mathbb{H} : \mathcal{M} \times \mathcal{A} \to \mathcal{M} \times \mathcal{A}$.

All the formal notation aside, the idea is that $\mathbb{H}$ is an operator that takes a matching μ and β in any particular round and produces a matching $\hat{\mu}$ along with the updated $\hat{\beta}$.

Figure 1:



That is, suppose we start from (μ^k, β). Then, $\mathbb{H}(\mu^k, \beta)$ is evaluated as follows.

3.3 Evaluation of $\mathbb{H}$

Remark 1. *An example of the evaluation of* $\mathbb{H}$ *follows after this subsection. It might be useful to go through the example before this subsection.*

Define $\mu(i) = 0 \forall i$.

Inputs are (μ^k, β).

Using the inputs, compute S^k making use of equations 3.3 and 3.4.

Define, $(\forall i \in M \cup W)$,

$$T_i^k := \{r : r \in S_i^k, i \in S_r^k\} \tag{3.5}$$

Define,

$$\alpha_i^k := \text{The } \succ_i \text{ maximum element in } T_i^k \tag{3.6}$$

STEP.1 Draw a graph $G^k = (N, E_k)$ where $N = M \cup W$, $(i, j) \in E_k$ iff $j = \alpha_i^k$. Naturally, if $T_i^k = \emptyset$ then, such nodes will not have any incoming or outgoing edges. Moreover, if a node has an incoming edge then it must have an outgoing edge too.

Definition 3.2. *The graph G^k is called Stage-1 Graph of $\mathbb{H}(S^k, \beta)$.*

STEP.2

Definition 3.3. *If there is a sequence of agents, $(a_1, a_2, .., a_n)$ $(n \geq 2)$ such that $a_{i+1} = \alpha_i^k$ (modulo n) then that constitutes a cycle.*

Let,

$$\mathcal{C} := \{C : C \text{ is a cycle in } G^k\} \tag{3.7}$$

Cutting at one edge determines all the matches in that cycle uniquely with matches (a_i, a_{i+1}) $(i = 1, 3, ..)$ or (a_i, a_{i-1}) $(i = 1, 3, ..)$ once again *modulo n*.

Break each cycle randomly at some edge.

That is, for a cycle $\{a_1, a_2, ..., a_n\}$ suppose we break it at a_1, a_2. Then, the match formed is -

$$\mu(a_2) = a_3, \mu(a_3) = a_2, \mu(a_4) = a_5, \mu(a_5) = a_4, \ldots \mu(a_n) = 1 \text{ and } \mu(a_1) = a_n \tag{3.8}$$

STEP.3 Every time we break a cycle a set of matches is formed. Remove all the nodes in those cycles and any edges pointing towards any of the nodes in cycles. Let the remaining set of agents be N^1.

Redraw the $G_k^1 = (N^1, E_k^1)$ as in Step 1 but only restricting attention to the remaining agents N_1.

- Define (in the context of graph G_k^1,

$$T_i^k := \{r : r \in N^1, r \in S_i^k, i \in S_r^k\} \tag{3.9}$$

$$\alpha_i^k := \text{The } \succ_i \text{ maximum element in } T_i^k \tag{3.10}$$

$$E_k^1 = E_k^1 \cup \{(i, j) : i \in N^1, j = \alpha_i^k\} \tag{3.11}$$

- Go to STEP 2.

ELSE proceed as below.

3.3.1 Updating μ^k and β

Define, for all i,

$$\hat{\beta}_i := \begin{cases} \min\{|A_i|, \beta_i + 1\} & \text{if } \mu(i) = 0 \\ \beta_i & \text{otherwise} \end{cases} \tag{3.12}$$

Recall that $|A_i|$ was the set of all acceptable agents according to i. The formulation above essentially prevents β_i from growing beyond the set of acceptable agents if the agent is unmatched. If the agent is matched, then β_i is not updated.

Update β and μ as below -

$$\beta_i = \hat{\beta}_i, \forall i \tag{3.13}$$

$$\mu^{k+1} = \mu \tag{3.14}$$

The operator $\mathbb{H}$ is evaluated as $\mathbb{H}(\mu^k, \beta) = (\mu^{k+1}, \hat{\beta})$.

3.3.2 Example of $\mathbb{H}$

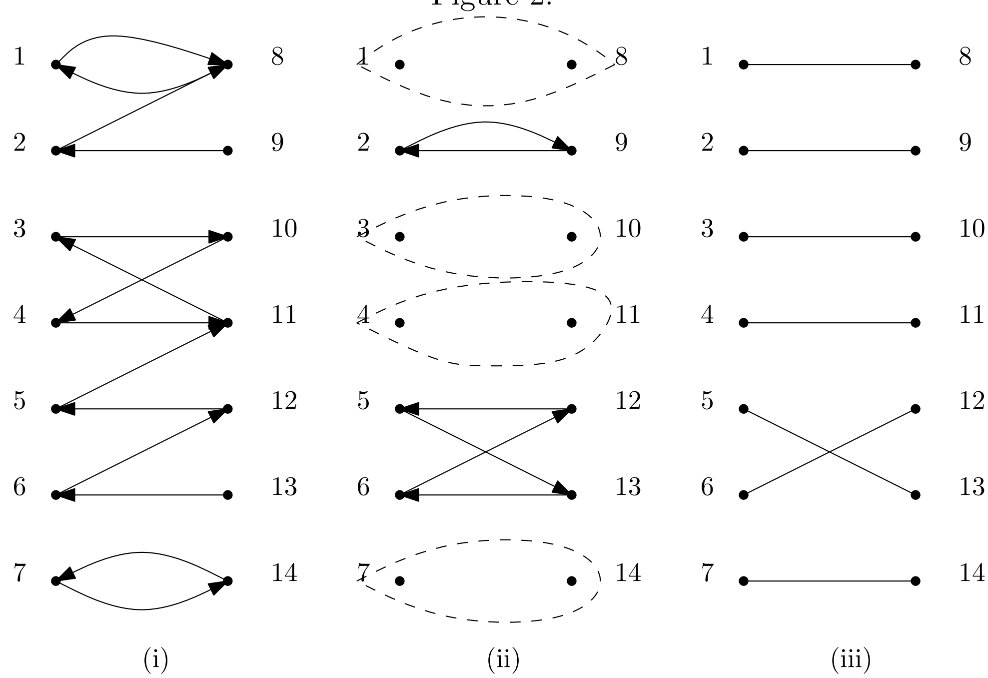
It would be better to illustrate the map $\mathbb{H}$ with an example of a Graph. Suppose we have 7 agents on either side, say $(1, 2, .., 7)$ are men and $(8, 9, ..., 14)$ are women. And let us suppose that in some intermediate round the we evaluate the S_i, T_i and α_i from equations 3.3, 3.5 and 3.6 to obtain the Stage-1 Graph. Suppose it looks like (i) in the figure below.

Then, first we eliminate all the cycles. We break cycles of length 4 or more randomly. By doing that we obtain a Graph as in (ii). The dotted agents are absent in that graph but have been shown only to show what matches were formed in the first stage. 1 finds 8 the best and 8 finds 1 the best amongst the set of agents they can potentially be matched to in this round. (This set is given by T_i). $3 \to 10 \to 4 \to 11 \to 3$ is a cycle that is broken randomly to generate a matching $(3, 10)$ and $(4, 11)$. $(2, 14)$ is another match formed as they are the most preferred agents according to each other.

Now, in the second stage, 2 can no longer point to 8 as 8 is matched in the first stage and is removed. So, in the graph in (ii), 2 points to, say, 9. Note that 2 might not point to any agents also. In this example, I have assumed that $9 \in S_2$ and $2 \in S_i$. That is, they find each other mutually acceptable *in this round*. Since this is a 2-cycle, it is broken uniquely to get a match $(2, 9)$. Similarly 5 can no longer point to 11 and so points to 13 now, giving a 4 cycle- $5 \to 13 \to 6 \to 12 \to 5$. This is broken randomly to generate a match, say, $(5, 13), (6, 12)$.

The end result of this round $\mathbb{H}$ is showin in the final graph in (iii). Note that the operation in this round $\mathbb{H}$ could result in some instability. For example, look at (i).

Figure 2:



Here, if 11's preferences are $3 \succ_1 15 \succ_1 14$ then there will be an instability. This is because when we broke the $4-$ cycle, 11 was matched to 4. Moreover, 5 pointing to 11 despite finding 12 acceptable implies that $11 \succ_5 12$. Therefore, in $(4, 11)$ and $(5, 12)$ is an unstable match. However, there is no attention paid to stability in each round. In the next round, $5 \in T_{11}$ and $11 \in T_5$ and they have a potential to match. Various other things could happen in the next round but the idea is a blocking pair will be *discovered* eventually.

We allow each cycle to be broken randomly *independently* of other cycles. Moreover, we also do not pay any attention towards stability.

3.4 Termination Condition

With the introduction of $\mathbb{H}$, in any round k, suppose the current matching is μ and the β is at β^k. Then, we evaluate $\mathbb{H}(\mu^k, \beta)$. Set the output equal to $(\mu^{k+1}, \hat{\beta})$.

That is, define,

$$(\mu^{k+1}, \hat{\beta}) := \mathbb{H}(\mu^k, \beta^k) \tag{3.15}$$

There is some abuse of notation here as μ^{k+1} is determined by breaking some cycles randomly and therefore for the same μ^k, β the outcome of $\mathbb{H}$ might differ in two different instances at some intermediate round. But, as will be proved later, with the termination condition, there will not be any randomness involved in the final round.

Naturally, we need a termination condition.

The termination condition is,

$$\text{STOP IF } \mu^{k+1} = \mu^k \text{ and } \hat{\beta} = \beta \text{ ,ELSE, GO TO ROUND } k+1 \tag{3.16}$$

Definition 3.4. *Call the pair (μ^k, β) the terminal input if the algorithm has terminated.*

If the algorithm does not terminate with the above termination condition, as it won't in many cases, then I propose a fix to terminate it. Although, I reserve it to section 5.2.

4 Examples

We look at a simple example that we saw in 3.1.2 first. Then, we look at a slightly more involved example. Lastly, we look at an example where the algorithm fails to terminate. We will come back to that example after we present the full proofs of the procedure to be followed when the algorithm does not terminate.

4.1 A simple example

Here, we present the preferences of the entire market in one table. Agents 1 to 4 can be thought of as men and 5 to 8 can be thought of as women.

1	2	3	4	5	6	7	8
5	6	7	8	2	3	4	1
6	7	8	5	3	4	1	2
7	8	5	6	4	1	2	3
8	5	6	7	1	2	3	4

The way to read this matrix is columns j denotes the preferences of agent j, ignoring the top row. That is, 1's preferences are $5 \succ_1 6 \succ_1 7 \succ_1 8$.

In the first two rounds nothing happens. There are no mutually proposing pairs and no matches are formed. That is,

$$T_i^k = \emptyset \ \forall i, \forall k \leq 2$$

In round 3,

$$T_1^3 = \{6.7\}, T_2^3 = \{7.8\}.T_3^3 = \{8.5\}, T_4^3 = \{5, 6\} \tag{4.1}$$
$$T_5^3 = \{3, 4\}, T_6^3 = \{4, 1\}, T_7^3 = \{1, 2\}, T_8^3 = \{2, 3\} \tag{4.2}$$

Therefore,

$$\alpha_1^3 = 6, \alpha_2^3 = 7, \alpha_3^3 = 8, \alpha_4^3 = 5 \tag{4.3}$$

and

$$\alpha_5^3 = 3, \alpha_6^3 = 4, \alpha_7^3 = 1, \alpha_8^3 = 2 \tag{4.4}$$

And so, we have a cycle -

$$1 \to 6 \to 4 \to 5 \to 3 \to 8 \to 2 \to 7 \to 1$$

Randomly cut the cycle at any point and you will either end up with

$$\mu_1 = \{(1,6),(2,7),(3,8),(4,5)\} \text{ or } \mu_2 = \{(1,7),(2,8),(3,5),(4,6)\}.$$

4.2 Example

Table 1: Men's Preferences

1	2	3	4	5
6	6	6	6	7
7	9	8	7	8
8	8	9	9	10
10	7	7	8	6
9	10	10	10	9

Table 2: Women's Preferences

6	7	8	9	10
4	5	2	4	1
5	4	1	3	2
1	1	3	2	3
2	2	4	5	4
3	3	5	1	5

- **Round 1:**

  $$T_i^1 = \emptyset \; \forall i \notin \{4,5,6,7\}$$

  .

  Observe that 4 proposes to 6 and 5 to 7 and vice-versa.

  Obviously,

  $$\mu^1 = \{(4,6),(5,7)\}$$

new Henceforth, let us *pretend* as if $(4, 5, 6, 7)$ do not exist as the match amongst themselves is not going to be broken.

That is, from now on on $N = \{1, 2, 3, 8, 9, 10\}$.

- **Round 2:**

$$T_i^2 = \emptyset \ \forall i$$

.

So no new matches are made in this round.

$$\mu^2 = \mu^1.$$

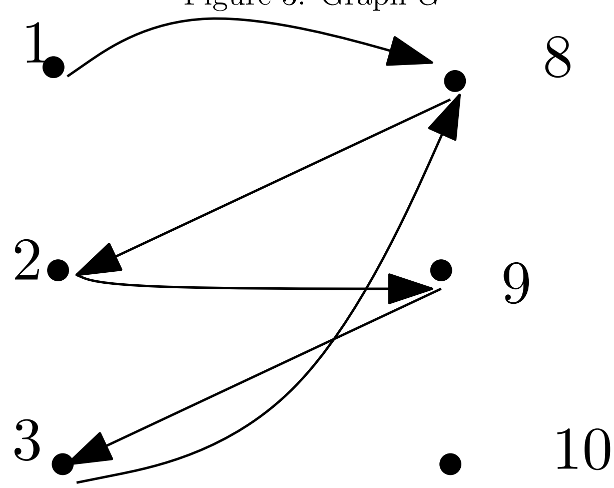
- **Round 3:**

$$T_1^3 = \{8\}, T_2^3 = \{8, 9\}, T_3^3 = \{8, 9\}, T_8^3 = \{1, 2, 3\}, T_9^3 = \{2, 3\}, T_{10}^3 = \emptyset$$

Note once again that all the S_i^3 are written without $4, 5, 6, 7$ to avoid unncessary cluttering.

Below is the graph of this round.

Figure 3: Graph G



The cycle from the diagram can be broken in two ways. Let us chose to break it favouring men.

So,

$$\mu^3 = \{(2, 9), (3, 8), (4, 6), (5, 7)$$

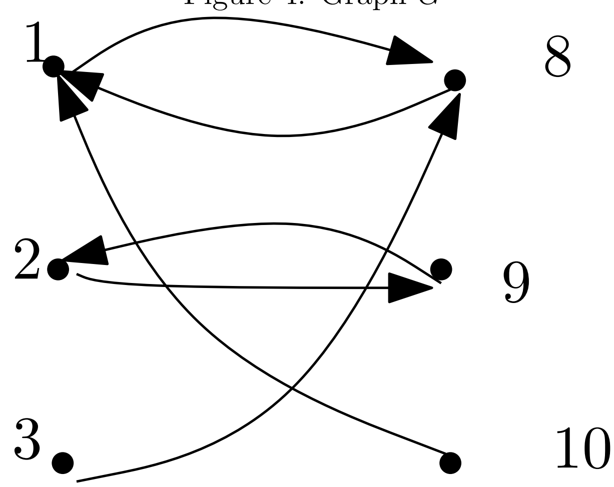
- **Round 4:**

  Now, 1 proposes to $S_1^4 = \{6, 7, 8, 10\}$. However, due to the matches made, the S_i^4 have shrunk as described in the algorithm.

  $$T_1^4 = \{8, 10\}, T_2^4 = \{9\}, T_3^4 = \{8\}, T_8^4 = \{1, 3\}, T_9^4 = \{2\}, T_{10}^4 = \{1\}$$

  The graph is as below -

  Figure 4: Graph G

  Once again, we eliminate the cycles first to obtain,

  $$\mu^4 = \{(1, 8), (2, 9), (4, 6), (5, 7)\}$$

- **Round 5:** 3 being kicked from his match with 8 in round 3, now proposes to $(6, 8, 9)$.

  $$T_1^5 = \{8\}, T_2^5 = \{9\}, T_3^5 = \{9\}, T_8^5 = \{1\}, T_9^5 = \{2, 3\}, T_{10}^5 = \emptyset$$

  The graph, as always, is below -

  This yields, after breaking the cycles,

  $$\mu^5 = \{(1, 8), (3, 9), (4, 6), (5, 7)\}.$$

- **Round 6:**

  Agent 2, after being kicked out of his existing match with 9 now expands his set S_2^4.

  $$T_1^6 = \{8\}, T_2^6 = \{8\}, T_3^4 = \{9\}, T_8^6 = \{1, 2\}, T_9^6 = \{3\}, T_{10}^6 = \emptyset$$

Figure 5: Graph G

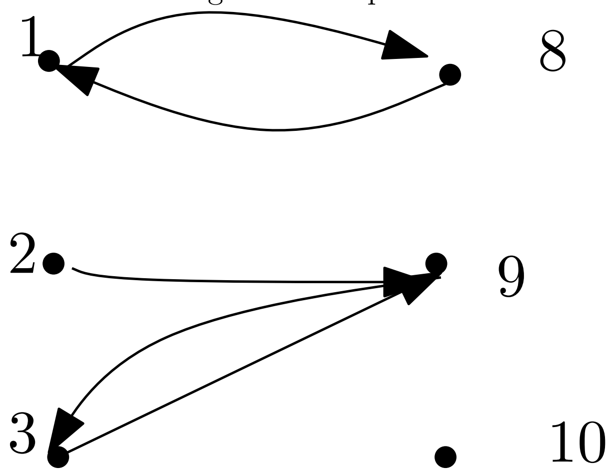
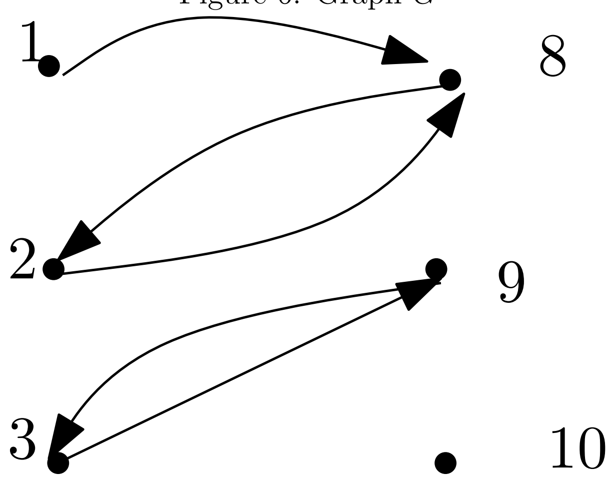


Figure 6: Graph G



And the graph -

Therefore,

$$\mu^6 = \{(2,8),(3,9),(4,6),(5,7)\}$$

- **Round 7:** Agent 1 was matched to 8 in previous round. So $S_1^6 = \{6,7,8\}$. After being kicked out at the end of round 6, 1 will expand his set and propose to 10 as well.

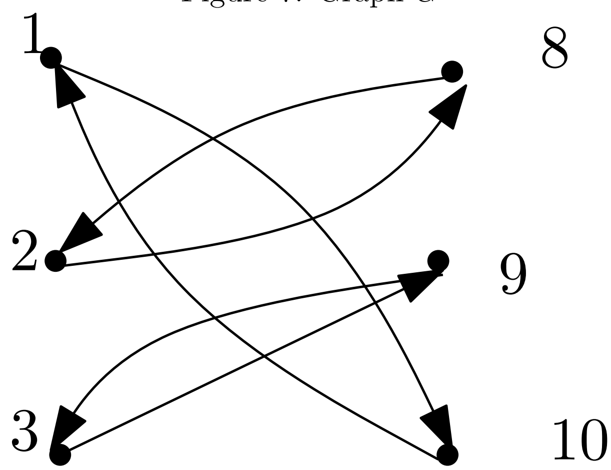
$$T_1^7 = \{10\}, T_2^7 = \{8\}, T_3^7 = \{9\}, T_8^7 = \{2\}, T_9^7 = \{3\}, T_{10}^{=}\{1\}$$

And therefore,

$$\mu := \mu^7 = \{(1,10),(2,8),(3,9),(4,6),(5,7)\}$$

Remark: A few remarks about this example are interesting. First of all, observe that μ obtained through the algorithm is also the Gale-Shapley outcome for Men proposing side. More importantly, this also coincides with the women proposing Gale-Shapley for this instance.

Figure 7: Graph G



4.3 An example where the algorithm fails to terminate

The preferences of men look as below -

1	2	3	4
6	7	8	5
7	8	7	8
5	5	5	6
8	6	6	7

And the preferences of women are -

5	6	7	8
4	3	1	2
1	4	4	4
3	1	2	3
2	2	3	1

Let us start with the algorithm. Observe that 4 and 5 appear on each other's top. Therefore, they will be matched in the first round and that match will never be broken.

Round 1 In the first round

$$T_i^1 = \emptyset, \forall i \in \{1, 2, 3, 6, 7, 8\}$$
$$T_4^1 = \{5\}, T_5^1 = \{4\}$$

Therefore,

$$\mu^1 = \{(4, 5)\}$$

Moreover,

$$\beta = (2, 2, 2, 1, 1, 2, 2, 2)$$

That is, $\beta_1 = \beta_2 = \beta_3 = \beta_6 = \beta_7 = \beta_8 = 2$. The reason is because these agents are unmatched in their first round and hence they increase their list by 1.

Round 2 Let us present a table for S_i, T_i and μ_i for each agent in this round. Recall that if an agent i is unmatched then $\mu(i) = 0$.

Table 3: Table of S_i^2, T_i^2, μ^2

Agent	S_i^2	T_i^2	μ^2
1	6,7	7	7
2	7,8	8	8
3	8,7	$\emptyset$	0
4	5	5	5
5	4	4	4
6	3,4	$\emptyset$	0
7	1,4	1	1
8	2,4	2	2

Therefore, at the end of this round, agents 3 and 6 are unmatched. So $\beta(3) = \beta(6) = 3$. β for all the other agents remains unchanged.

Round 3 Let us look at a similar table as above in this round. We also add β to this table instead of presenting it at the end of the round. It should be kept in mind that the β in the table corresponds to the updated β at the end of this round.

Table 4: Table of S_i^3, T_i^3, μ^3 and β

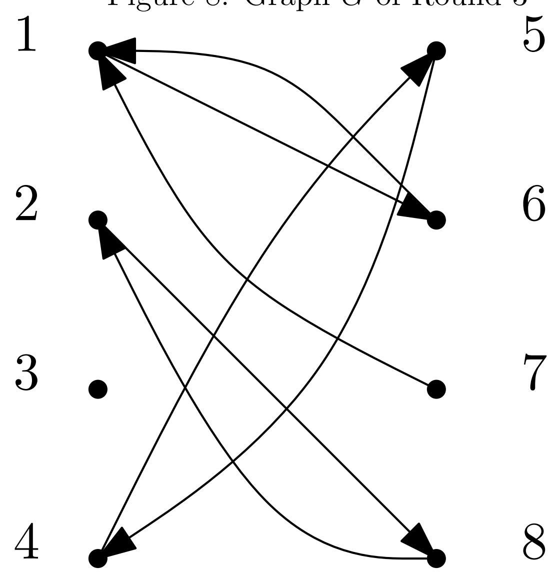
Agent	S_i^3	T_i^3	μ^3	β
1	6,7	6,7	6	2
2	7,8	8	8	2
3	8,7,5	$\emptyset$	0	4
4	5	5	5	1
5	4	4	4	1
6	3,4,1	1	1	3
7	1,4	1	0	3
8	2	2	2	2

To see why the match in this round has changed, to be precise 1 has left 7, his match from Round 2 for 6, look at the Stage-1 Graph of this round. Since 1 is matched to 7 in Round 2, 1 and 7 are in each other's T_i in this round. But in this round, 6 proposes to 1 and 1 likes 6 better than 7. So in the graph 1 points towards 7.

At the end of this round, therefore, 7 is unmatched and its β is updated to 3.

$$\mu^3 = \{(1, 6), (2, 8), (4, 5)\}$$

Figure 8: Graph G of Round 3



Stage Graph of Round 3

Round 4 Continuing as before, here is the table for Round 4.

Table 5: Table of S_i^4, T_i^4, μ^4 and β

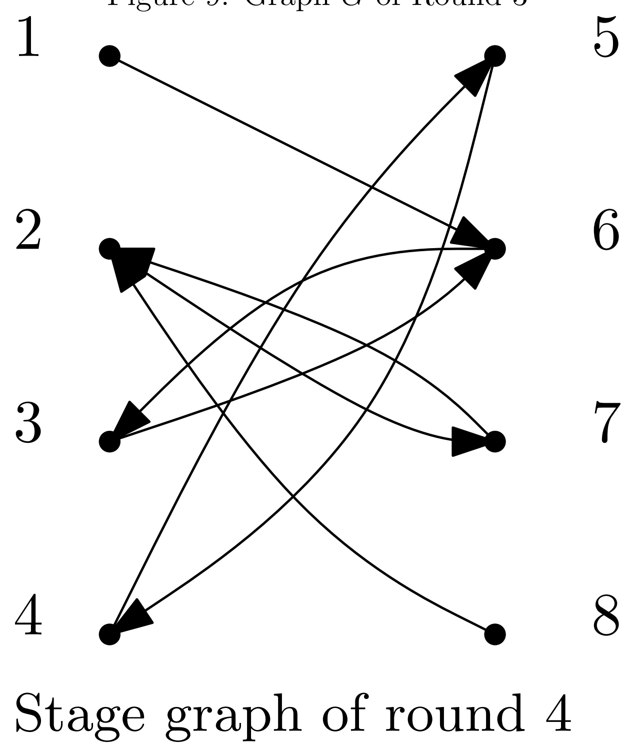
Agent	S_i^4	T_i^4	μ^4	β
1	6	6	0	3
2	7,8	7,8	7	2
3	8,7,5,6	6	6	4
4	5	5	5	1
5	4	4	4	1
6	3,4,1	1,3	3	3
7	1,4,2	2	2	3
8	2	2	0	3

Below is the graph for this round.

Since very soon we will start the cycling phase, it is important to note what μ^4.

$$\mu^4 = \{(2,7),(3,6),(4,5)\}$$

Figure 9: Graph G of Round 3



Round 5 And we continue the process.

Table 6: Table of S_i^5, T_i^5, μ^5 and β

Agent	S_i^5	T_i^5	μ^5	β
1	6,7,5	7	7	3
2	7	7	0	3
3	8,7,5,6	6,8	8	4
4	5	5	5	1
5	4	4	4	1
6	3	3	0	4
7	1,4,2	1,2	1	3
8	2,4,3	3	3	3

Therefore,

$$\mu^5 = \{(1,7),(3,8),(4,5)\}$$

From now on, it can be easily checked that the algorithm keeps cycling through the following three matchings without termination.

$$\gamma_1 := \{(1,7),(3,8),(4,5)\} \tag{4.5}$$
$$\gamma_2 := \{(1,6),(2,8),(4,5)\} \tag{4.6}$$
$$\gamma_3 := \{(2,7),(3,6),(4,5)\} \tag{4.7}$$

For those interested, you can jump to 5.5 to see how this problem is resolved for this example. The proofs for a general case will be in the next section.

5 Results

Remark 2. *For any graph obtained in the above construction, if a node has an incoming edge then it must have an outgoing edge.*

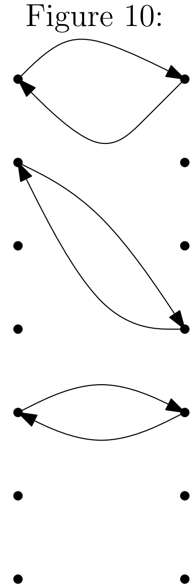
Remark 3. *If, for any node i, $deg^-(i) > 0$ then $deg^+(i) = 1$.*

Remark 4. *For any i, $deg^+(i) \leq 1$.*

Definition 5.1. *For any Stage-1 graph (Defn. 3.2), call it of Type-1 if it consists of only 2-cycles and singletons.*

For example, Figure 10 on the next page is of Type-1.

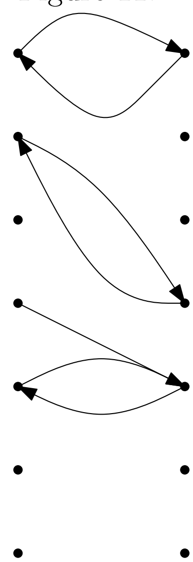
Figure 10:



While the Figure 11 on the next page is NOT of Type-1

Lemma 5.1. *In any round k, if the Stage-1 Graph of $\mathbb{H}(\mu^k, \beta^k)$ is not of Type-1 then Round k cannot be a terminal round.*

Figure 11:



Proof. In other words, what we need to prove here is that if the Graph is not of Type-1 then for some agent i, either S_i^k is updated or β_i is updated. We will prove that S_i^k is updated.

The idea is rather simple. Suppose $G = (N, E)$ is not of Type-1. $\Rightarrow E$ (the set of edges) is non-empty. Since our bipartite graph is such that if $deg^-(i) > 0 \Rightarrow deg^+(i) = 1$, set of edges being non-empty means the only possible configurations are cycles of even length with or without incoming edges to them. Since our assumption is that G is not of Type-1, $\exists$ a cycle $C = \{a_1, ., a_n\}$ with or without some incoming edges to some nodes.

Case 1: Length of $C > 2$.

Our algorithm entails breaking this cycle randomly. Let's assume we break it at a_1, a_2. That is, we match (a_1, a_n) and (a_2, a_3). Since $a_1 \to a_2 \to a_3 \to ...a_n \to a_1$, we have that $a_3 \succ_{a_2} a_1$.

Since this is a terminal round, it must be the case that $\mu^k = \mu^{k+1}$. Therefore, we must have $\mu^k(a_1) = a_n$ and $\mu^k(a_2) = a_3$. Since the cycle is $a_1 \to a_2 \to a_3 \ldots \to a_n$, we have that $a_3 \succ_{a_2} a_1$. But then, since we had $\mu^k(a_2) = a_3$ to begin with, $a_1 \notin S_{a_2}^k$. Therefore, edge (a_1, a_2) is not possible.

Therefore, we cannot have a cycle of length > 2.

Case 2: Length of $C = 2$, $\forall C$.

Given that our assumption is that G is not of Type-1, this means that $\exists C$ of length 2, say (a_1, a_2) with an edge, say $a_3 \to a_1$.

Now, if we $\mu^k(a_1) = a_3$, then at the end of this round, (since a 2-cycle will be broken in a unique way) we will have $\mu^{k+1}(a_1) = a_2$. Therefore, this cannot be a terminal round.

If $\mu^k(a_1) = a_2$ then, $a_2 \succ_{a_1} a_3$. Therefore, edge (a_3, a_1) is not possible in this stage.

That is, if there is a cycle of length 2 then it cannot have any incoming edge from outside.

∎

Lemma 5.1 can be restated as, *If a round k is a terminal round, then the Stage-1 Graph of it must have only two-cycles and isolated points.*

Proposition 5.1. *If the algorithm terminates, then the matching μ obtained in the last round is stable.*

Proof. Suppose not.

Say, round k is the terminal round. Let the round input be (μ^k, β^k).

Therefore, if
$$(\mu^{k+1}, \beta^{k+1}) = \mathbb{H}(\mu^k, \beta^k)$$
then, $\mu^{k+1} = \mu^k$.

Suppose, $\exists$ a blocking pair $(i.j)$. So, $i \succ_j \mu^k(j)$ and $j \succ_i \mu^k(i)$. Since $j \in T_i^k$ and $i \in T_j^k$, $deg^+(i) = deg^+(j) = 1$.

By lemma 5.1, we know that the stage graph of $\mathbb{H}$ applied on the terminal round's μ^k, β^k would be of Type-1.

That is, $\exists$ an agent m such that (i, m) form a two-cycle in the Stage-1 Graph. Moreover, since Stage-1 Graph is drawn by each agent pointing to its top agent in T_i^k, i cannot be pointing to $\mu^k(i)$ as $j \in T_i^k$ and $j \succ_i \mu^k(i)$. By Type-1 nature of the graph, i will be matched to $m \succ_i \mu^(i)$.

$\Rightarrow \mu^{k+1}(i) \neq \mu^k(i)$.

A contradiction.

∎

For the sake of Lemma to appear next, consider any stable matching μ. Define β as follows -

$$\beta_i = \begin{cases} \text{Rank of agent } \mu(i) \text{ according to } \succ_i & \text{if } \mu(i) > 0 \\ |A_i| & \text{if } \mu(i) = 0 \end{cases} \tag{5.1}$$

Recall that A_i was the set of all the acceptable agents for i.

Lemma 5.2. $\mathbb{H}(\mu, \beta) = (\mu, \beta)$*, where μ is some stable matching and β is defined accordingly as per the equation above.*

Remark 5. *That is, Lemma 5.2 says that if in a particular round, the algorithm encounters a stable matching such that all the unmatched agents are at the end of their proposal sets, then the algorithm terminates at that matching.*

Proof. Since μ is stable, $\nexists$ a pair (i, j) such that $j \succ_i \mu(i)$ and $i \succ_j \mu(j)$.

Therefore, for any $j \in \{k : k \succ_i \mu(i)\}$, $\mu(j) \succ_j i$.

Therefore, $j \notin T_i$.

So,

$$T_i = \begin{cases} \mu(i) & \text{if } \mu(i) > 0 \\ \emptyset & \text{if } \mu(i) = 0 \end{cases}$$

So, the stage-1 graph will be of Type-1. We will break the cycles uniquely with $\hat{\mu(i)} = \mu(i)\ \forall i$ such that $\mu(i) > 0$.

Moreover, since for the unmatched agents, $\beta(i) = |A_i|$, the β_i is not updated for any agent.

Therefore, $\hat{\mu} = \mu$.

So, $\mathbb{H}(\mu, \beta) = (\mu, \beta)$. $\square$

5.1 Termination of the algorithm

Suppose the algorithm does not terminate. Then it must cycle since there are only finite number of possible matchings. Moreover, in every round of the cycling either the β increases for those who are unmatched or it stays constant if they are matched or the β_i has reached its maximum. Thereofore, it is without loss of generality to say that throughout the cycle, β remains constant.

Definition 5.2. *A matching μ is stable restricting attention to set S if $\nexists$ a blocking pair (i, j) such that $i, j \in S$.*

5.2 Resolution of a Cycle

We need an appealing procedure to correct the problem of cycling. The problem of cycling arises due to the fact that the algorithm is trying to satisfy all the agents in one round. That creates potential instabilities in subsequent rounds. However, intuitively, even though the algorithm cycles, it eliminates several blocking pairs on the way and therefore all we need to do is pick any existing matching in the cycle and fix it.

The procedure is a slight modification of the *Random Order Mechanism* proposed by Ma (1996) [8] which was based on Roth and Vande-Vate (1990)[11]. In their random priority mechanism they start with an empty matching and pick an agent randomly to add to the existing matching. Here, since the algorithm is cycling it already has an existing matching. The matching itself may have blocking pairs. So we first eliminate the blocking pairs in a particular way and then start with the procedure. The procedure is as in the following subsection.

5.2.1 Modified map $\hat{\mathbb{H}}$

Before we begin the presentation, let us modify the map $\mathbb{H}$ from subsection 3.3 slightly. A quick recap of what we do in the map $\mathbb{H}(\mu, \beta)$ given a μ, β.

For each i (this is important), we construct S_i as

$$S_i = \{j : j \succ_i \mu(i)\}$$

Then, we construct T_i as

$$T_i = \{j : j \in S_i, i \in S_j$$

Then we draw a stage graph where each agent points to his(her) top agent from T_i. Following this we break cycles randomly and construct a graph with the remaining agents until we reach an empty graph.

In the modified map $\hat{\mathbb{H}}(\mu, \beta)$ the only change we bring is we ignore the agents that are unmatched in μ and have $\beta_i = |A_i|$. That is, we set their $S_i = \emptyset$ for all those agents who have exhausted their entire list of acceptable agents and are unmatched. Following this they will never be matched if we repeatedly apply $\hat{\mathbb{H}}$.

The following lemma says that if we repeatedly apply $\hat{\mathbb{H}}$ on any matching μ, we will converge in finite time. The lemma is stated with β as maximum but that is rather trivial as, if that is not the case then for the unmatched agents the β will increase in each round until the β becomes maximum.

Let $\hat{\mathbb{H}}^n(\mu, \beta)$ denote applying $\hat{\mathbb{H}}$ on (μ, β) n times.

Lemma 5.3. *Consider $\beta_i = |A_i|$ for each i. That is, β is maximum. $\hat{\mathbb{H}}^n(\mu, \beta)$ converges to a matching in at most $|M| + |W| + 2|M||W|$ rounds for any matching μ.*

Proof. Suppose all the agents are matched in μ. If it is stable then we have immediate convergence. Suppose not.

In any round at least one of the following three cases must happen -

- Some agents go from matched to unmatched and hence are *out of the market.*
- Some or all agents strictly improve upon their match.
- (μ, β) remain unchanged.

Firstly, if in some round, $\hat{\mathbb{H}}(\mu, \beta) = (\mu, \beta)$ then, as proved in 5.1, the Graph is of Type-1 and hence this round is terminal. If the round is not terminal then one of the first two cases above happens.

There are at most $|M| + |W|$ agents to be unmatched and at most $2|M||W|$ possible improvements possible. Therefore, after $|M| + |W| + 2|M||W|$ rounds, $\hat{\mathbb{H}}$ has no scope of either of the first two possibilities. Therefore, the matching must remain unchanged. (Since β is assumed to be maximum for the lemma. This is without loss of generality as if it is not, there are $2|M||W|$ possible improvements of β.

Moreover, similar to Lemma 5.1, if $\hat{\mathbb{H}}(\mu, \beta) = (\mu, \beta)$ remains unchanged then the Graph obtained must be of Type-I. That is, there are no blocking pairs amongst the set of matched agents in the matching obtained through repeated application of $\hat{\mathbb{H}}$. $\square$

Therefore, we obtain the initial matching for the *Random Priority Mechanism* as below -

- Choose any matching in a cycle. Call the corresponding matching μ.
- Evaluate $\mathbb{H}^n(\mu, \beta)$ until it converges.
- Call this matching μ_0 and start with the Random Priority Procedure due to Ma (1996) as below.

This is the procedure to obtain a stable matching starting from the match μ_0 obtained from above. (Produced from Klaus and Klijn (2006) [5]

- **Input:** Matching μ_0 obtained as above and (M, W, P).

  Set $R_0 = \{i : \mu_0(i) > 0\}$. Let $\hat{n} := |N \backslash R_0|$. Set $t = 1$.

- **Step t:** Choose an agent i_t from $N \backslash R_{t-1}$ at random. Set $R_t = R_{t-1} \cup \} i_t \}$. Suppose $i_t = w \in W$. (Otherwise replace w by m in Step t.)

- **Stable Room Procedure:**

  Case 1: $\nexists$ a blocking pair (m, w) in μ_{t-1} with $m \in R_{t-1}$. Stop if $t = \hat{n}$ and define output $\mu = \mu_{t-1}$. Otherwise, set $\mu_t = \mu_{t-1}$ and go to step $t := t + 1$.

  Case 2: $\exists$ a blocking pair (m, w) for μ_{t-1} with $m \in R_t$. Choose the blocking pair (m^*, w) for μ_{t-1} with $m^* \in R_t$, that w prefers the most.

  If $\mu_{t-1}(m^*) = m^*$, then define μ_t such that $\mu_t(w) := m^*, \mu_t(m^*) := w$, and for all $i \in N \backslash \{w, m^*\}$, $\mu_t(i) := \mu_{t-1}(i)$. Stop if $t = \hat{n}$ and set the output of the algorithm $\mu := \mu_t$. Otherwise go to step $t := t + 1$.

  If $\mu_{t-1}(m^*) = w'$, then redefine $\mu_{t-1}(w) := m^*$, $\mu_{t-1}(m^*) := w$, $\mu_{t-1}(w') := w'$, and for all $i \in N \backslash \{m^*, w, w'\}$, $\mu_{t-1}(i) := \mu_{t-1}(i)$. Set $w = w'$ and repeat the **Stable Room Procedure**

While the above description is from Klaus and Klijn (2006) [5], a description of the same procedure using the Graph-theoretic structure might be more transparent.

To give a graph theoretic formulation of the above procedure, like the one we have been using, let $\mathbb{H}_S(\mu, \beta)$ denote execution of $\mathbb{H}$ from subsection 3.3 restricting attention to the agents in S. That is, we run $\mathbb{H}$ as if agents outside S do not exist.

5.2.2 Random Priority Mechanism Using Graph

- **Input:** The matching μ_0 obtained through repeated application of $\hat{\mathbb{H}}$ and (M, W, P). Set $R_0 = \{i : \mu_0(i) > 0\}$.

  Observe that $\mathbb{H}_{R_0}(\mu_0, \beta) = (\mu_0, \beta_0)$ and the corresponding Stage Graph is of Type-1.

  Let $\hat{n} := |N \backslash R_0|$. Set $t = 1$.

- **Step t:** If $t > \hat{n}$, define output $\mu = \mu_{t-1}$ and STOP. Otherwise, choose an agent i_t from $N \backslash R_{t-1}$ at random. Set $R_t = R_{t-1} \cup \} i_t \}$. Suppose $i_t = w \in W$. (Otherwise replace w by m in Step t.)

- **Stable Room Procedure:** Execute $\mathbb{H}_{R_t}(\mu_{t-1}, \beta_{t-1})$ repeatedly until it converges to some (μ_t, β_t). (To be proved below that this always happens.) Go to step $t = t + 1$.

Lemma 5.4. *The repeated application of $\mathbb{H}_{R_t}$ starting from (μ_{t-1}, β) converges.*

Proof. The proof will be using induction. First, let's prove the claim for $t = 1$. Suppose $i_1 = w$. Only two things can happen upon adding w to the set R_0. Either there is a blocking pair or not. Since μ_0 had no blocking pair restricting attention to R_0, any blocking pair must involve w. So, Stage-1 graph of of $\mathbb{H}$, if there is a blocking pair would as one of the two pictures below.

In either case, the graph will only have two cycles and upon breaking them, every man is weakly better off. At least one man is strictly better off due to getting matched to w. And since w gets matched to some man m, $m's$ earlier match, if any, is now unmatched. In every iteration of $\mathbb{H}$, at most one woman goes from matched to unmatched. In this process a man she was matched to initially strictly improves. If no woman goes from matched to unmatched then either an unmatched woman gets matched to a single man or a woman remains unmatched. In either case, we have convergence. If one man strictly improves in every round, this can happen only finite number of times and hence will stop in at most $|M||W|$ rounds.

Therefore, starting from initial μ_0, we obtain μ_1 that is stable restricting attention to R_1.

Since the above argument only used the fact that μ_0 was stable restricting attention to R_0, in any round t, if μ_{t-1} is stable restricting attention to R_{t-1} then we must obtain a matching μ_t that is stable restricting attention to R_t. □

5.3 Algorithm In Full Form

We can present the above *"fix"* to the algorithm to present the entire algorithm as below -

1. Start the algorithm as described in Section 3.

2. If it cycles, pick any matching μ^* from the cycle. Obtain a matching μ_0 through repeated application of $\hat{\mathbb{H}}$ on the existing μ^*, β until we obtain a matching μ_0.

3. Use the Random Priority Mechanism *starting with* μ_0 as described in 5.2.2.

We can put all the above points in a theorem.

Theorem 5.2. *The algorithm, as presented in 5.3 produces a stable matching for every preference profile P.*

Proof. This is obvious as, if the algorithm does not cycle it yields a stable matching according to Proposition 5.1. If it cycles, then we use the *Random Priority Procedure* starting with one of the matches in a cycle and turn it into a stable match *restriction attention to the matched agents.* (Recall Defn. 5.2.) Then at every stage, the outcome of RP mechanism is stable restricting attention to the matched agents up to that round. Since in the last round, there are no blocking pairs the outcome is stable. □

5.4 Return to the cycling example with a fix

Recall the Example 4.3. The algorithm there cycled through the following three matchings.

$$\gamma_1 := \{(1,7),(3,8),(4,5)\} \tag{5.2}$$
$$\gamma_2 := \{(1,6),(2,8),(4,5)\} \tag{5.3}$$
$$\gamma_3 := \{(2,7),(3,6),(4,5)\} \tag{5.4}$$

Observe that any of $\gamma_1, \gamma_2, \gamma_3$ are stable restricting attention to the matched agents in the corresponding matches.

Let us start with γ_1 as the initial matching. The unmatched agents are $2, 6$. Below is the Random Priority Mechanism starting from γ_1.

- Start with $\mu_0 := \gamma_1$. $R_0 = \{1,3,4,5,7,8\}$.
- We pick one of $2, 6$ randomly. Suppose we pick 6. The only blocking pair involving 6 is $(1,6)$. So we match $(1,6)$ Therefore, $\mu_1 := \{(1,6),(3,8),(4,5)(7)\}$.
- Now we pick 2. The two blocking pairs are $(2,7)$ and $(2,7)$. But 2 prefers 7 and hence we match them to obtain, $\mu_2 := \{(1,6),(2,7),(3,8),(4,5)\}$.

Note that had we picked 2 instead of 6 in the first round, we would have found a different matching. Both the matchings are listed below.

$$\mu_M = \{(1,6),(2,7),(3,8),(4,5)\} \tag{5.5}$$
$$\mu_W = \{(1,7),(2,8),(3,6),(4,5)\} \tag{5.6}$$

An important concept to be seen through this example possibly suggesting why we need a *Random Priority Procedure* is because when we looked at both 2 and 6 that we were unmatched in γ_1 in one round, they disturbed the matching of 1 and 8 simultaneously. 1 was matched to 7 who became unmatched due to arrival of 6. And 2 became matched to 8, at the expense of 3. In the subsequent round, 7 arrives and matches with 2, leaving 8 unmatched. Similarly, 3 arrives and matches with 6, leaving 1 unmatched. The entire problem, is every new agents is coming with a *lag* of one round causing this cycling. Dealing them one by one takes care of this problem.

6 Fairness Properties

Part of the motivation for the algorithm in the introduction was due to a *fair procedure*. As mentioned before, various randomized procedures such as those in Klaus and Klijn (2006) [5] (Employment by Lotto proposed by Aldershof et al [1] or Random Order Mechanism by Ma [8] that is based on Roth and Vande-Vate) put a positive probability on every stable matching irrespective of the number of stable matchings and the preference profiles. In the two examples before the formal presentation of the algorithm, the median stable matching by Sethuraman et.al. [12] would pick the matchings that lie in the center in the Hasse diagram of the poset of the stable matchings. However, it was also proved that obtaining a median stable matching is a computationally hard problem.

One possible and perhaps necessary approach towards studying fairness in the context of stable matchings is axiomatic. The two most natural axioms that can be motivated purely from fairness standpoint would be *Gender Neutrality* and *Permutation Invariance.* That is, any *fair* procedure should not distinguish between Men and Women and should also be immune to relabeling men or women. As Gokturk and Masarani (1989)[9] showed that there does not exist a deterministic stable matching mechanism that satisfies these two properties. Therefore, even restricting ourselves to these two seemingly necessary requirements for fairness we ought to have randomized mechanisms. One possible randomized mechanism could be to have a uniform lottery on the entire set of stable matchings. Another mechanism could be to select one of the two Gale-Shapley outcomes with equal probability. Both these mechanisms would satisfy *Gender Neutrality* and *Permutation Invariance.*

In this regard, it can be easily noted that even the algorithm proposed in this paper would satisfy both these properties. Since every time we encounter a cycle of length 4 or more, it is broken randomly, the correspondence of the set stable matching that this algorithm could potentially obtain would be unchanged if interchange the role of men or women or relabel any men or women. Therefore, this algorithm could be viewed as a new procedure to produce a set of stable matchings. From the perspective of fairness, this would be the main contribution of the algorithm. As mentioned before, besides fairness, the other appealing property of the algorithm is that it has a decentralized nature and does not need a clearinghouse.

7 Conclusion

In this article I focus on developing a new procedure that does not distinguish between men and women. I propose a new procedure that has an appealing feature of both sides actively proposing in each round. The most commonly used procedure in two-sided matching markets is Gale-Shapley (1962)'s Deferred Acceptance Algorithm. The DAA has only one side proposing in each round and the other side accepting or rejecting proposals. In contrast, here in each round, both sides of the market make proposals and only mutually acceptable agents are matched in each round.

The algorithm has an appealing property that it does not distinguish between the two sexes, ex-ante. Moreover, the algorithm is also permutation invariant. This procedure should be viewed as an addition to the existing procedures that satisfy these two axioms. In the quest to obtain stable matchings that treat both the sexes equally, Sethuraman and Teo (2006) proposed the *Median Stable Matching.* A practical problem with the *Median Stable Matching* is we need to compute the entire set of stable matchings in order to obtain the *Median Stable Matching.* Secondly, and perhaps more importantly, there are no known axioms that the *Median Stable Matching* rule is known to satisfy. In contrast with the other proposed rules satisfying Gender Neturality and Permutation invariance (Klaus and Klijn, (2006), Romero-Medina (2002)) this algorithm does not require an ordering over agents to start the algorithm. Both the procedures need an explicit priority order over agents in order to begin the algorithm.

A natural extension of the procedure is to characterize the matching it obtains. Through various examples, it is seen that the algorithm often selects outcomes that are different from the DAA outcomes. Despite some randomness involved in breaking the cycles, this procedure is different from selecting a random stable matching or a random matching and eliminating blocking pairs proposed by Roth and Vande (1990).

Thus, this algorithm, should primarily be viewed as an attempt to address a remark by Knuth mentioned in the introduction. In contrast with the other procedures that address this concern, by imposing an ordering over agents, this is the first procedure wherein both sides participate in each round, beyond just accepting or rejecting offers.

References

[1] Aldershof B., Carducci O.M. and Lorence D.C., *Refined inequalities for stable marriage*, Constraints, 4:281-292, 1999.

[2] Echenique F. and Yariv L., *An experimental study of decentralized matching*, Caltech Social Science working paper, 2012.

[3] Gale D. and Shapley L.S., *College admissions and the stability of marriage*. American Math. Monthly, 69: 9-15, 1962.

[4] Irving R., Leather P. and Gusfeld D., *An efficient algorithm for the optimal stable marriage*. Journal on Computing, 16:111-128, 1987.

[5] Klaus B. and Klijn F., *Procedurally fair and stable matching*, Economic Theory, 27:431-447, 2006

[6] Klaus B. and Klijn F., *Median stable matching for college admissions*, International Journal of Game Theory, 34:1-11, 2006

[7] Knuth D., *Marriages Stables*, Montreal: Les Presses de l Universite de MOntreal, 2,3, 1976.

[8] Ma J. *On randomized matching mechanisms*, Economic Theory, 8:377-381, 1996.

[9] Masarani F. and Gokturk S.S, *On the existence of fair matching algorithms*. Theory and Decisions, 26: 305-322, 1989

[10] Roth A.E. and Sotomayor M.A.O., *Two-sided matching: A study in game-theoretic modeling and analysis*, Econometric Society Monograph Series, New York: Cambridge University Press, 1990.

[11] Roth A.E. and Vande Vate J.H. *Random paths to stability in Two-Sided matching*, Econometrica, 58:1475-1480, 1990.

[12] Sethuraman J., Teo C.P. and Qian L., *Many-to-one stable matching; Geometry and Fairness*, Mathematics of Operations Research, 31(3):581-596,2006.

[13] Teo C.P. and Sethuraman J., *The geometry of fractional stable matchings and its applications*, Mathematics of Operations Research 23:874-891,1998.